Information and Public Service Provision: Experimental Evidence from School WASH Services in Bangladesh Pre-Analysis Plan

Umrbek Allakulov[1], Serena Cocciolo[2], Binayak Das[1], Ahasan Habib[3], and Anna Tompsett*[4]

[1]Water Integrity Network
[2]Institute for International Economic Studies
[3]NGO Forum for Public Health
[4]Stockholm University

September 25, 2019

Abstract

This document describes the design and analysis plan for an impact evaluation of a participatory information-gathering and dissemination intervention applied to school WASH (Water, Sanitation and Hygiene) provision in rural Bangladeshi schools. The intervention was implemented in 30 schools in two upazilas in Bangladesh between September and December 2018. This document outlines the hypotheses to be tested and the specifications to be used in the empirical analysis. This document was written during follow-up data collection. None of the follow up data was examined or analyzed before finalizing this report. We expect to conduct additional exploratory analyses. When reporting results, we will indicate which analyses were pre-specified and which are exploratory.

*Correspondence: anna.tompsett@ne.su.se.
This project is realized in partnership with the Water Integrity Network (WIN), NGO Forum for Public Health, and Development Organisation of the Rural Poor (DORP). AEA RCT registration: AEARCTR-0003111.

1 Motivation

Effective public service provision is critical for human welfare but often undermined by corruption or inefficiency. Many policymakers claim that providing information to stakeholders can help improve public service delivery, by making service providers more accountable to service users. However, previous evidence is inconclusive. Some studies do find that providing information to beneficiary communities can indeed improve public service delivery (Reinikka and Svensson, 2005). Others find no effect of information without complementary interventions to directly incentivize stakeholder action (Banerjee et al., 2010a; Björkman et al., 2017). Our study contributes to this policy debate by evaluating the Annotated Water Integrity Scan (AWIS) intervention, a scalable, replicable, and adaptable information-gathering and dissemination tool.

AWIS is a participatory approach to information gathering and dissemination designed to evaluate integrity, defined as a set of practices that impede corruption and promote respect for the rule of law. AWIS is also specifically designed to initiate discussions that lead to the identification of priority steps for improvement. AWIS was developed by the Water Integrity Network (WIN) as a portable tool that can be easily adapted to different contexts and intervention areas. AWIS has previously been applied in countries including Benin, India, Uganda, and Guatemala.

In this study, we evaluate AWIS in the context of WASH (Water, Sanitation and Hygiene) provision in rural Bangladeshi schools. Many schoolchildren in Bangladesh lack access to adequate WASH facilities. Lack of access to adequate WASH facilities increases the prevalence of water-borne disease and absenteeism in school-age children and may increase infant mortality in younger siblings (UNICEF, 2012). In Bangladesh, the average school has only one toilet for every 187 students, and only a quarter of toilets are clean (Water Aid, 2016).

Worldwide, corruption is a major obstacle to adequate WASH service provision (UNDP, 2011). In school WASH services in particular, funds for building or maintaining toilets may be siphoned off. As a result, interventions to increase stakeholder awareness about legal responsibilities, funding allocations, and channels for complaints can create incentives for improved provision.

In order to measure the impact of AWIS on WASH provision in schools, we implement a randomized controlled trial (RCT) in which we randomly assign schools to a treatment group, which receives the AWIS intervention, and a control group, which does not. We then collect systematic data before and after the intervention, including an inventory of school sanitation facilities, automated anonymized tracking of latrine use over a 72-hour period, student attendance and exam results. Our design allow us to evaluate whether a single, time-limited information intervention can improve WASH service provision and education outcomes in villages in rural Bangladesh.

Our results will help contribute to the policy debate regarding how best to improve public

service provision. Despite the mixed prior evidence, citizen engagement continues to be promoted as an approach to improving public service delivery, and international aid agencies increasingly condition access to their funds on the adoption of beneficiary participation components (Banerjee et al., 2010b). Our results will help to understand whether such requirements are warranted.

2 Context

We work in 60 high schools[1] in Bhola Sadar and Ramgati upazilas in southern Bangladesh. Both are classified as "Hard to Reach" areas, a classification the Bangladeshi government uses to describe areas with poor water and sanitation coverage and particular obstacles to improvement.[2] Since school sanitation facilities are most important for female teenage students, we work only in schools attended by both female and male students (49 schools) or by female students only (11 schools). In our study schools, baseline student enrolment was on average 576 students with a mean student:teacher ratio of 46.5 students per teacher, corresponding to 12 teachers per school. This is slightly worse than the country average of 34 students per teacher.[3] In total, 34,547 students were enrolled in the 60 study schools, of which 20,217 were female.

Our local partner NGO identified a pool of schools to target for the intervention, out of which we randomly assigned schools to treatment and control. The local partner NGO identified local schools with sanitation facilities below national standards. In practice, almost all schools in the local areas were selected for the study. No school declined to participate in the study and we did not exclude any schools from the pool selected by the partner NGO for any other reason.

Our baseline data confirms that sanitation facilities in our study sample do not meet Bangladeshi and international standards, either in terms of quantity or quality. Enumerators identified on average 4.5 toilet cubicles available to students, of which on average 3 were functional. The ratio of students to toilets was thus on average 218:1, worse than the 187:1 country average and far below the internationally recommended standard of 50:1 (Water Aid, 2016). Among functioning toilet cubicles accessible to students, 91% had a door that could be closed and 54% were clean. Only 9% were accessible for disabled students. The majority of toilet cubicles lacked features included in the minimum requirements for safe and private menstrual hygiene management in schools UNESCO (2014), for example 57% did not have running water inside and 79% did not have a bin for disposal of menstrual products.

School sanitation facilities are important determinants of female adolescent student at-

[1] High schools cover grades 6 to 10.

[2] Specifically: unfavourable hydrogeological conditions, inadequate communications networks, and/or frequent natural disasters.

[3] Country-level data is the latest available data (2017) extracted from the UNESCO Institute for Statistics.

tendance. Among female students interviewed, only 19% reported that school attendance is unaffected by menstruation: 8% reported that girls do not usually come to school when they have their period and a further 73% that girls leave school earlier when they have their periods. The vast majority of girls interviewed (85%) cited poor or inadequate school sanitation facilities as one of the reasons why some girls miss school when they have their periods. Among the reasons for the inadequacy of the school sanitation facilities, female students cited lack of private facilities for girls to change (64%), to wash cloth pads (51%), and to dispose of sanitary napkins (34%).

Reports from students contrast with levels of awareness among school staff and management about the needs of female students with respect to sanitation facilities for Menstrual Hygiene Management (MHM). For example, when asked about these needs, only 37% of headteachers cited separate toilet cubicles for boys and girls, 23% a mechanism for the collection and disposal of sanitary napkins and 18% a space for girls to privately wash cloths. Awareness of student needs regarding MHM appears higher among female teachers than male teachers.[4] However, with only two female headteachers in our sample and with fewer than 20% of teachers and school management committee (SMC) members female, this additional awareness appears unlikely to be captured in school decision-making processes. In this context, the AWIS intervention may help to disseminate information about best practices for hygiene and MHM management and to involve all stakeholders in decision-making, especially female teachers and female students.

3 Intervention

The cornerstone of the AWIS approach is a participatory workshop conducted with local stakeholders. The AWIS workshop is a single, time-limited intervention that can reveal new information by facilitating constructive dialogue between stakeholders. The process may help policy-makers, local government officials, civil society organisations, and user representatives to establish priority actions to enhance water integrity and governance. The AWIS process is designed to allow discussion of integrity without antagonizing stakeholders and to help raise awareness of contrasting points of view and unexpected risks.

In this study, we implement AWIS with respect to WASH (Water, Sanitation and Hygiene) service provision in schools in rural Bangladesh. In order to implement AWIS in this context, WIN adapted the approach to suit the context, including with respect to national

[4]For example, when asked about the characteristics that a sanitation facility should have for safe MHM, a much higher share of female teachers than male teachers cited the importance of a mechanism for the collection and disposal of sanitary protection materials (56% vs 12%) or of a space to privately wash cloths (62% vs 16%). When asked about possible ways to make it easier for girls to manage their periods while they are at school, female teachers' responses very closely resembled those of female students, as both groups largely cited the importance of private facilities to dispose of sanitary napkins (88% vs 83%), to wash (62% vs 66%) and to wash cloths (56% vs 54%).

policy e.g. the Bangladeshi government's 11-point guidelines for school WASH provision. However, the AWIS process is adaptable to other contexts and scales.

The AWIS workshop is divided in three main stages. During the first stage, stakeholders, including students, parents, school staff and local government administrators, anonymously score the quality of WASH services and WASH-related processes in the school on indicators for transparency, accountability and participation. Participants complete an anonymous questionnaire (scorecard) covering 5 key areas: (i) quality of sanitation facilities, (ii) gender, (iii) menstrual hygiene, (iv) disability, and (v) budget and expenditure (Appendix B).

During the second stage of the workshop, participants are invited to openly discuss each aspect included in the scoring process. Participants are asked to share their motivations for high and low scores for each question in the scorecard. The focus of this stage is not on defending individual opinions but on understanding why different participants gave different scores. Based on the outcomes of the discussion, the scores are jointly adjusted and aggregated, until they are agreed upon.

Finally, each participant identifies priority areas for action and specific actions that can be taken. The results of this exercise are openly disclosed to all participants, and local facilitators ensure that the priority areas and proposed actions are clearly understood by all participants.

The implementing NGO carried out the AWIS intervention between September and December 2018 (see Figure 1 for the full project timeline). Project staff complete a detailed report after each workshop, including participant information, workshop processes, discussion outcomes, and the main lessons learned.

The number of participants in the AWIS workshops ranges between 18 and 21. Participants mostly comprised parents, students, teachers and local leaders, of which the majority were informal, social leaders rather than formal representatives of local administration (Table 1). Among participants, 36% were female. More female students and parents participated than female teachers or SMC members. Almost all students participating (97%) were between 14 and 16 years old.

Participants scored each indicator on a scale of 1 to 3. Workshop participants were least satisfied with the provision of sanitation facilities for disabled students and with facilities for Menstrual Hygiene Management (MHM) (Table 2). The scoring system is imperfect. Although scoring is intended to be anonymous, some participants required assistance to complete the scoring process. Some participants may not have fully understood all the questions. Scores were recorded anonymously, so we are cannot evaluate whether participants with different characteristics gave different scores. However, in general, scores given by participants correlate positively with direct, independent observations by enumerators during school visits (Appendix Table H1), suggesting that the information captured by the AWIS process is meaningful.

Table 3 and 4 report, respectively, the most cited priorities and actions identified by the

participants to the AWIS workshop.[5] On average, participants at each workshop identified 20 priority areas and 20 actions. The most commonly identified priority areas included improvements to physical infrastructure. In contrast, the most commonly proposed actions were relatively "soft": seeking support from local administration, holding meetings, or increasing awareness among students.

Workshop participants evaluated the strengths and weaknesses of the AWIS workshop during the workshop itself. Perceived strengths included the educational content of the workshop with respect to best WASH and MHM practices, and its interactive and discursive nature (Tables 5). Participant critiques of the process were often practical, for example reporting that the time allocated to the workshop was too short. Participants felt that multimedia or practical examples would have facilitated understanding. Also, many participants perceived the isolated nature of the AWIS intervention as a weakness, citing a lack of follow-up or ongoing activities as a weakness. Participants also reported that the process needed to be wider in scale, including more people (Table 6).

4 Randomization of the intervention

We randomly assigned 30 schools to receive the intervention and 30 schools to a control group who will receive no intervention via this project.[6] We assigned schools to treatment and control at two public lotteries, one held in each of the study upazilas. Assignment to treatment is thus stratified by upazila, with 15 schools in each upazila assigned to treatment and 15 to control. We collected baseline data before assignment to treatment and control. All schools had the same ex-ante probability of receiving the program.

We report balance tests on school characteristics and characteristics of sanitation facilities in Tables 7 and 8 respectively. The results show that baseline differences between schools assigned to treatment and control are consistent with differences that could arise due to chance.

We invited school and local representatives to the public lottery meetings to ensure full transparency of the randomization process. We motivated the presence of the control group by explaining that the budget for the intervention is limited, and that including the control group allows us to measure the impact of the program, which in turn provides information to improve policy in the future.

[5]Participants listed priorities and actions descriptively. Research assistants categorized responses into discrete categories.

[6]Importantly, we do not prevent them from receiving interventions through any other program.

5 Data collection

We carried out a baseline survey in all study schools between July and August 2018, before assigning schools to treatment and control, and before implementing the project in treated schools. We will conduct the follow-up survey between August and October 2019, approximately 8-12 months after the implementation of the AWIS workshop in treated schools and 13-14 months after collecting baseline data. We will use these data to evaluate program impacts. Key outcomes include indicators of WASH (Water, Sanitation and Hygiene) provision, school attendance, test results on standardized national exams, and anonymous measures of latrine use intensity, captured using motion sensors and event loggers.

Documentation of the AWIS procedure itself also provides rich information about how the process was implemented, how stakeholders respond, and the problems faced by stakeholders in improving WASH provision. We discuss these data in Section 3 and we expect to further exploit these data in exploratory analyses after testing our main hypotheses.

5.1 Baseline survey

We conducted interviews in all schools with staff and students. We interviewed the headteacher regarding school programs and initiatives related to WASH and MHM (Menstrual Health Management); student, parent and teacher involvement in WASH facility management; and the school budget and institutional processes, especially with respect to WASH facilities.[7] We interviewed several teachers in each school regarding the content of school curricula related to WASH and MHM; the system of responsibilities within the school for cleaning of WASH facilities; and student awareness of WASH and MHM practices. We interviewed school cleaning staff regarding the availability of cleaning products and the cleaning/maintenance schedule. We interviewed one male and one female student regarding whether and how WASH and MHM practices are discussed and taught in school. We also asked headteachers, teachers and students to report on the quality of the existing WASH facilities in the school; on the constraints faced by female students during menstruation and their effect on school attendance; and on priorities to improve WASH services and practices in the school.

Importantly, during the interviews we did not ask questions about private or sensitive data. In particular, we did not ask questions about individual WASH or MHM practices. We phrased all questions related to WASH and MHM practices in general terms, asking respondents to answer questions based on their own and their peers' experiences.

We also collected data from all schools on WASH facilities, student attendance, exam results, and latrine use. Enumerators observed and photographed all WASH facilities in

[7] In Bangladesh the headteacher is always a member of the School Management Committee (SMC) by law. In our study schools, all headteachers are appointed as secretary of the SMC.

study schools and recorded whether they complied with basic WASH and MHM standards. WASH facilities include drinking water points, toilet cubicles, urinals, or handwashing stations. We test all sources of drinking water available in the school for bacteria and arsenic contamination. We collect data regarding compliance withfor each school sanitation facility and its components.

We obtained attendance data from school attendance registers and by head-counts performed by enumerators during the school visit. To conduct head-counts unobtrusively, enumerators asked students to list improvements they would like to make in the school and vote for and against each suggestion, recording the total number of students and the number of students in favor of each suggestion. We collected audit data to verify and quantify the extent of misreporting in school registries. We were particularly concerned that attendance of female students might be over-reported, given incentives created by a stipend program designed to encourage female student participation.[8] In practice, we find limited evidence of manipulation, with the exception of male students in Grade 10, for whom official records appear to overstate attendance (Appendix Table H2).

We downloaded individual examination results for study schools from official records publicly available online. We collect individual test results on the most recent standardized national exams for grade 8 (Junior School Certificate, JSC) and grade 10 (Secondary School Certificate, SSC). Results for both JSC and SSC exams are publicly available online on the education board's web-based result publication system. We recorded individual GPAs and student gender.[9]

We used event loggers to obtain latrine use data in a discrete and non-intrusive manner. Whenever latrines have functional doors, we use magnetic sensors that record opening and closure of latrine doors. Alternatively, we use motion sensors, which record when a warm body passes through its detecting area. Both sensors were small and packaged to render them unobtrusive. We use unique codes in order to identify sensors and link sensor data with the corresponding school-toilet cubicle. Enumerators installed sensors after school hours or during class time, in order to interfere as less as possible with the normal use of bathrooms by students. Enumerators removed the sensors three days after installation, excluding Fridays and holidays.

[8]The Female Secondary School Stipend Project pays tuition fees and provides monthly stipends for unmarried rural girls up to grade 10 who attend recognized institutions, remain unmarried, maintain at least 75% attendance, and secure at least 45% marks in annual examinations (a pass requires 35%).

[9]In practice, this is a two-step process. School-level records list individual student results using student ids. Detailed individual student data is also publicly available, searchable by student id. We first recorded the exam results for each school and then determined the gender of each student with a recorded result. Gender is not recorded on the official record but research assistants determined gender using student names, which are mostly gender-specific in Bangladesh.

5.2 Follow-up survey

At follow-up, we will adopt the same data collection instruments as at baseline, ensuring that our data are comparable across time. In addition, we will add a module on the AWIS intervention in treated schools, asking about knowledge of the intervention, participation in the intervention, and opinions about the process.

We expect minimal attrition of schools from the study. School participation in the follow-up survey is conditional on informed oral consent from the headteacher. The headteacher of all schools included in our final sample already consented to baseline data collection. We randomly assigned schools to treatment and control at public lottery meetings with school representatives, guaranteeing transparency and legitimacy of the process. Our implementing partner reported generally positive reception of the AWIS workshop among participants. For these reasons, we do not foresee cases where the school agreed to participate in the study at baseline but not at follow-up.

6 Sample

Our sample consists of 60 high schools, equally divided between two upazilas, Bhola Sadar and Ramgati. Recruitment was finalized at baseline during school visits, when we obtained informed oral consent from the headteacher to conduct the study in the school (Appendix A). Headteachers from all targeted schools gave consent to conduct the study, although in a few cases headteachers declined consent for parts of the study.

6.1 Interviews

We obtain informed consent to conduct the study in the school from the headteacher and from each subject before conducting the interviews (Appendix A.1, A.2). Within each school we interview headteachers, teachers, students, and cleaning staff. We expect turnover among staff and, especially, students, so we do not attempt to create a panel of respondents. Our sample is therefore a repeated cross-section.

We design our sampling protocols to target respondents using a consistent approach in both treatment and control groups at both baseline and follow-up. At follow-up, we also augment our sampling approach in treated schools so that we additionally interview students and teachers that participated in the AWIS workshops, if no students and teachers that participated in AWIS are selected as the main respondents. Table 9 describes our approaches to sampling among each targeted group of respondents.

Baseline interview response rates We successfully interviewed the headteacher in all schools (60 headteachers).

We randomly selected one male and one female teacher for interview from the full roster of teachers who were present at the schools during the survey visit. In 11 schools, there was no female teacher present at school during our school visit, so we targeted a total of 60 male teachers and 49 female teachers. Additionally, we selected for interview 128 male teachers and 45 female teachers with special responsibility for WASH (Water, Sanitation and Hygiene) and/or MHM (Menstrual Hygiene Management).

We successfully interviewed at least one male and at least one female teacher at all schools where teachers of the given gender were present, but we did not successfully interview all targeted teachers.[10] In 53 cases (42 female teachers and 11 male teachers), the teacher selected for the interview was not interviewed either because of a shortage of time or because of a miscommunication with enumerators regarding how to complete the survey form. In another 53 cases (all male teachers), the teacher did not give consent to the interview.[11] There is no systematic difference in sample selection between treated and control schools (Appendix Table H3). Teachers who were listed later in the roster were less likely to be successfully interviewed (Appendix Table H4). At follow-up, we revised the sampling protocol, reduced the targeted number of teachers, clarified the directions with the field staff, and modified the survey constraints to ensure compliance with the revised protocol.

.

We selected a total of 109 students for interview, comprising one male and one female student in each of the schools with students of the given gender.[12] We completed the interview with 108 of them, 106 in grade 10 and 2 in grade 9. One male student declined to participate in the survey. The mean age of student interviewees was 15.2 years. We interviewed one 13-year-old and eleven 14-year-olds. The oldest student interviewed was 17.

Only 8 schools reported having any cleaning staff, only 5 of which were present on the day of the survey. We interviewed these 5 janitors.

6.2 Observation of WASH facilities

At baseline, we recorded and tested 60 sources of drinking water, one in each school. We observed a total of 331 sanitation facilities. Among the 331 sanitation facilities, 267 were available for students while 65 were reserved for teachers' use, and 132 were for males only, 128 for females only, and 71 for both genders. Within these facilities, we recorded data on 305 toilets, 123 urinals, and 350 sinks.

[10] Of the randomly-selected teachers, we successfully interviewed 49 male teachers and 35 female teachers. Among the teachers selected because of responsibility for WASH or MHM, we successfully surveyed 75 male teachers and 17 female teachers.

[11] No female teacher did not consent to be surveyed. Among male teachers, 28% of targeted teachers did not consent to the interview.

[12] We identified more than one eligible male student in 6 schools and more than one eligible female student in 17 schools. In these cases, we randomly selected one student of each gender.

6.3 School attendance

We collect attendance data from school records for the month when the survey takes place and additionally for February and March, two months in which no major religious festivals or national examinations interrupted school attendance. We obtain audit attendance data in each grade, disaggregated by gender, based on enumerator head counts.[13] In each school, we collect official and audit attendance data for one classroom per grade, randomly sampling one classroom in case the school has more than one classroom per grade.

At baseline, we collected attendance data in 58 schools. In two schools, the headteachers declined to give consent for us to collect attendance data, either from school registries or audit data. In the 58 schools where we obtained consent, we identified 290 classes for data collection.[14]

We collected attendance data from the schools registries for 290 classes for the month of the school visit (July or August 2018), 289 classes for March 2018 and 286 classes for February 2018.[15] Our baseline sample consists of 22,215 students for the month of the school visit (July or August 2018), 22,109 students for March 2018 and 21,066 students for February 2018.

We collected audit data from 289 classes.[16]

6.4 Exam results

We collect exam results data for all 60 study schools. Our baseline sample consists of 10,847 students, of which 6,430 took the 2017 JSC exam and 4,417 took the 2017 SSC exam. On average, we have exam result data for 107 students per school for the JSC exam and 74 students per school for the SSC exam. About half the students taking exams were female: 54% of the students writing the JSC exam and 49% of the students writing the SSC exam.

The number of students for which we have exam results is larger than the number of registered students, because students from smaller schools cross-register at larger schools to sit exams. This is a potential attenuating factor in our evaluation of effects on exam results, because some students whose exam results are listed for our study schools will have attended different schools and thus not been exposed to the AWIS intervention. The ratio of exam results to registered students in the respective grades is 1.3:1 in treated schools and 1.2:1 in control schools.[17] We will report whether this relationship remains stable over time in our

[13] We designed the survey instrument such that, in case of more than one class per grade, we would collect audit data on attendance for one randomly extracted class.

[14] In 79% of cases there was only one classroom per grade, therefore we randomly selected one classroom for the data collection for 61 grades in 20 schools.

[15] The cases of missing data are due to technical errors when attaching the pictures from the class registry to the survey form.

[16] We lost data for one class because of a temporary technical error with the tablet.

[17] The difference is not statistically significant.

follow-up data.

6.5 Latrine use

We collect data on latrine use for a maximum of 4 toilet cubicles per school. We only collect monitoring data for toilets that are functional and available to students. If the school has more than four cubicles that are both functioning and available to students, we randomly select which 4 toilet cubicles to monitor. For toilets that are not functional, we code the use as zero.

At baseline, we identified a total of 179 toilet cubicles that were eligible for sensor installation across all study schools. We randomly selected 145 of these for usage monitoring. We did not install sensors in three toilet cubicles. In one school, the headteacher did not consent to sensor installation (one toilet cubicle). For two toilet cubicles, we did not have any sensor available for installation at the time of the school visit. In total, we installed 124 magnetic sensors and 18 motion sensors. Sensors were installed for an average of 3 days.[18] We lost data from 8 sensors because enumerators did not correctly upload the log files from the sensors[19] and from 6 sensors for which we have partial missing data because they stopped functioning during installation.[20]

6.6 Power calculations

All schools assigned to the treatment group decided to receive the intervention and participate in the AWIS workshop, therefore we are able to use the full sample of 60 schools for the impact evaluation in most analyses. Our most conservative power calculations do not account for the panel data structure (one observation before and one observation after the intervention) and exploit only the variation across schools. These calculations suggest that the study is powered to detect minimum detectable effects (MDEs) (with significance level 10% and power 80%) of 0.55 standard deviations. These MDEs correspond to relatively large effects. Using baseline and panel data will increase power as long as the school and time fixed effects absorb more than half of the variation (Burlig et al., 2017; McKenzie, 2012), which we expect them to do for most of our outcome variables, especially those that correspond to physical infrastructure.

Many of our analyses will also be powered to detect smaller effects as a result of obtaining multiple observations within each school at baseline and follow-up. We do not know of any analytical expressions to account for both multiple observations over time and clustered

[18]Excluding 3 schools where the sensors were installed for 16-18 days due to a long break at the end of August, the average duration of the sensor installation was 3 days and never less than 2 days. There was no incidence of stolen or lost sensors.

[19]In two cases, the log file is missing. In six cases, the log file does not match the installation dates.

[20]We drop data for sensors if the enumerators reported that the sensors were broken or non-functional when they removed the sensors, and if we have fewer than 100 observations in the raw data.

treatment status. Accounting for the number of students for whom we have exam results data, without accounting for the panel data structure, reduces the MDEs for effects on exam results to 0.33 standard deviations.[21]

7 Hypotheses and variable definition

7.1 Main hypotheses

The theory of change underlying our study design is shown in Figure 2. The AWIS intervention is designed to spread information about priority areas for action and potential courses of action, and to strengthen institutions. In turn, these changes should improve WASH (Water, Sanitation and Hygiene) service provision and may impact on WASH facility use patterns. Improved WASH service provision may in turn increase school attendance and translate into improved exam performance. Motivated by this theory of change, we evaluate the effect of the AWIS intervention on the following three families of outcomes: knowledge, processes and institutions (A); WASH provision and use (B); and educational outcomes (C).

Throughout, we follow Anderson (2008) and use inverse covariance weighting to combine several sub-indices in a single summary index. This approach weights information from several standardized outcomes by the inverse of their covariance matrix in order to maximize the amount of information captured in the index. Following the original formulation in Anderson (2008), we calculate the index across all non-missing data for each observation separately, ensuring that we retain the maximum possible amount of information in construction of the index.

A Knowledge, processes and institutions We measure knowledge of WASH and MHM (Menstrual Hygiene Management) standards using an index for knowledge among head-teachers (A1), teachers (A2), and students (A3). Each index is created from several sub-indices measuring the share of correct responses given to a number of questions about best practice for WASH and MHM. We list the full set of sub-indices in Appendix C.

We measure institutional quality using an index capturing the presence and quality of institutions to manage WASH facilities and provide support to students for MHM (A4). We list the full set of sub-indices in Appendix D. We note that all the sub-indices we use to create the institutional quality index rely on self-reported data and are thus subject to concerns about experimenter demand effects and reporting bias. If we find effects only on outcome A4, we will interpret those results with caution.

[21] We use the analytical expression given in Djimeu and Houndolo (2016) and the observed baseline intra-cluster correlation of 0.37.

B WASH provision and use We measure WASH provision and use via three indicators.[22] First, we measure the number of functional toilet cubicles available for student use (B1), where we define a toilet cubicle as an individual stall, seat, squat-plate or drop-hole where a single person can defecate in private, and we define toilet cubicles as functional if they can be used and are not broken, damaged, or full.[23] Second, we create an index which measures the average quality of WASH facilities available to students (B2), including functionality, cleanliness, and the presence of items required for hygienic practices and for MHM. We list the full set of sub-indices in Appendix E. Third, we count the number of "toilet use events" per day in toilets available to students (B3) using data from door and motion sensors. We describe our current approach to construction of the count data from the raw sensor data in Appendix F.[24]

C Educational outcomes We measure educational outcomes using two metrics, school attendance (C1) and exam results (C2). We measure school attendance using a dummy variable for whether each student was present on a given day, using official registries for February and March by grade. We measure exam performance using student GPA[25] from the national Junior School Certificate (JSC, grade 8) and Secondary School Certificate (SSC, grade 10) exams.

In each case, the null hypothesis we test is that the intervention does not have a positive effect on the outcome, where positive is defined as greater knowledge, more comprehensive or active institutions, better provision of sanitation facilities in terms of quantity and/or quality, higher student attendance, or better exam results. Table 10 shows balance tests for all the main outcomes as measured at baseline.[26].

7.2 Secondary hypotheses

WASH provision is of particular importance for girls. We will therefore estimate the impact of the AWIS intervention separately for female and male students where it is possible to separately measure outcomes by gender. These outcomes are: student knowledge, number of functioning toilets available for each gender, quality of toilets available for each gender,

[22] We note that at baseline, before assignment to the intervention, 4 treated schools and 4 control schools reported that one sanitation facility was currently under construction and scheduled for completion by the end of 2018.

[23] We do not include urinals.

[24] We may be able to improve our classification algorithms once we have access to the follow-up data. If so, we will clearly identify and motivate any changes made to these algorithms in the main text, and compare results using our original approach to results using the improved approach.

[25] We code failing students, who receive a mean GPA of less than 1 but for whom the specific GPA is not reported, as having GPA equal to 0.5.

[26] Appendix Tables H5 to H9 provide additional balance tests on subindices.

latrine use, school attendance, and exam results.[27] We will report outcomes separately for male and female students, and we will additionally test in a pooled regression whether we can reject the null hypothesis that effects are equal for male and female students.

7.3 Exploratory analysis

We expect to carry out additional exploratory analyses which we will use to interpret the results of the pre-specified analyses. The primary goal of our exploratory analyses will be to understand the channels that mediate any effect we find, or, in case of a failure to reject the null hypothesis, to learn about why the intervention did not yield measurable effects on the pre-specified outcomes. We will clearly identify all exploratory analyses in the final report.

In exploratory analyses, we may describe actions taken by the school management committee to improve school WASH facilities, changes in the content of the school budget and management procedures, and changes in whether or not hygiene and MHM practices are taught in school. Our data will permit us to evaluate differences between treatment and control schools.

Our data will also allow us to describe awareness and understanding of the AWIS process, how AWIS participants themselves describe the AWIS process, whether the participants took any specific actions following the AWIS workshop, whether any such actions were successful in obtaining their objectives, and the perceived reasons for success or lack of success. We may also be able to evaluate whether the areas identified for priority action during the AWIS intervention show particular improvement between baseline and follow-up.

We will also be able to compare teachers that did and did not participate in the AWIS workshop in treated schools with teachers in the control schools and thereby evaluate whether we find evidence for knowledge spillovers. We note that we will be cautious in interpreting these results given that we cannot observe which teachers would have participated in the AWIS workshop in control schools.

We can also further explore other dimensions of our rich dataset, for example using the photographs we take of sanitation facilities directly to classify toilet cleanliness, either by classifying the images manually or by training a classifier on a subset of the images.

We also expect to organize focus group discussions (FGDs) with key groups of stakeholders. These FGDs will provide complementary qualitative evidence which will help us interpret our quantitative results. With the primary goal of our exploratory analysis in mind, we will finalize the design of the the FGDs in the light of the results of the tests of our primary hypotheses.

[27] For toilet quality and latrine use we will distinguish between sanitation facilities only for girls, only for boys, or for both.

8 Empirical analysis

We test our hypotheses using the following specification, which treats all data as a repeated cross-section:

$$y_{ist} = \alpha_s + \delta_{jt} + \beta \left(T_s \times Post_t \right) + \beta_u \left(T_s \times Post_t \times U_j \right) + \epsilon_{ist} \tag{1}$$

where y_{isjt} is an outcome variable in unit of observation i in school s in upazila j at time t. The indicator T_s takes the value one if school s is assigned to the AWIS intervention and zero otherwise. The indicator $Post_t$ takes the value zero before treatment and one after treatment. We include school fixed effects, α_s and upazila-specific time fixed effects, δ_{jt}, which control for baseline differences across schools and trends that affect both treated and control schools equally. In some specifications, we augment the school and time fixed effects to exploit other dimensions of our data, as we discuss below.

We allow effects to vary across upazilas by including the term $T_s \times Post_t \times U_j$, where U_j is an upazila control which takes the value 0.5 in Bhola and -0.5 in Ramgati. This approach ensures that the main coefficient of interest β correctly estimates the average difference-in-differences between treated and control villages, accounting for stratification in assignment to treatment (Gibbons et al., 2019; Imbens and Rubin, 2015; Lin, 2013).

We will test our main hypotheses using one-sided tests on the main coefficient of interest β. We use one-sided tests because the direction of our main hypotheses is clear and because this maximizes power to detect effects of the expected sign. We do not anticipate perverse effects (effects of the opposite sign to the hypothesised effects). However, they are hypothetically possible. For example, the AWIS intervention could discourage female students from attending school during menstruation if they become more sensitized to the inadequacy of school WASH (Water, Sanitation and Hygiene) facilities but AWIS is not successful in improving WASH service provision. If we estimate perverse effects for our primary hypotheses, we will additionally report whether or not they would have been rejected using a two-sided test, so that future research can incorporate this information in research design, with the caveat that the two-sided tests were not pre-specified and that those findings must thus be considered exploratory.

The specific command we will use to estimate coefficients is `reghdfe`, absorbing the fixed effects. The unit of observation i and the fixed effects we include vary according to the outcome variable, as shown in Table 11. To ensure that each school counts equally in summary statistics in all analyses, we weight observations by the inverse of the number of observations per school.[28]

[28]For outcome A2, the weights account for differential probabilities of being included in the sample for different types of teacher and specifically for oversampling of teachers with special responsibility for WASH or MHM. For outcome B3, the weights account for differential probabilities of inclusion in the sample for functioning and non-functioning toilets. For outcomes C1 and C2, we weight observations so that each grade

As treatment is assigned at the school level, we cluster by school when we report standard errors and confidence intervals. Because the number of schools involved in the study is small (30 per arm of the study), we conduct inference by randomization-based inference (RBI). Specifically, we reshuffle treatment status at each treatment assignment lottery 1,000 times and estimate the distribution of $\hat{\beta}$ under the null hypothesis of no positive treatment effect. We compare each estimated $\hat{\beta}$ under the observed treatment assignment to the distribution of the $\hat{\beta}$ under the null to generate p values.

Multiple hypothesis testing We test a total of nine primary outcomes in three families. There is thus a relatively high likelihood that we reject the null hypothesis in at least one of these tests due to chance, even if the null hypothesis is in fact correct. We will report three p values for each of our main primary outcomes: naïve p values, which would be correct if we had tested any one of the primary outcomes in isolation; p values that control the family wise error rate within each of our three outcome "families"; and finally p values that control the family wise error rate across the full set of nine primary outcomes. To construct these p values, we will follow Westfall and Young (1993), accounting for correlation between outcomes.

Robustness As a robustness test, we will estimate the main effects using control variables, using the Lasso algorithm to select the optimal set of controls i.e. the set of controls that minimizes the standard errors of the estimated treatment effect.[29]

We will test whether school enrolment or the number of students taking exams changes differentially in treatment and control schools. This will allow us to correctly interpret our results on attendance and exam results. Additionally, we will report whether the relationship between officially recorded and observed attendance changes between baseline and follow-up and whether treatment affects these changes. These tests will allow us to evaluate whether changes in reporting behaviour affect our results.

For results using exam results, we will report additional results using a dummy for whether or not the student passed the exam. Additionally, we will test sensitivity to different approaches to coding GPA for failing students.

We will also examine alternative measures of latrine use intensity. First, we will report an alternative measure of latrine use intensity, the total time per day in which the latrine appears to be in use. This measure addresses a potential problem with the motion sensor data, which undercounts events during periods of consecutive toilet events. Second, we will aggregate the sensor data to the school level, by calculating the mean number of toilet events per toilet cubicle, multiplying by the number of toilet cubicles, and dividing by the number

(C1) or exam type (C2) contributes equally, implying that the average does not change if the distribution of students across grades or exam types changes between baseline and follow-up.

[29] We report in Appendix G the list of controls included in the Lasso algorithm.

of students present at school. This measure will allow us to correctly interpret the latrine use data in the presence of changing numbers of functional toilets or changing numbers of students. When we aggregate the latrine use data to the school level, we will report results in which we measure latrine use both using the event count data and using the total event time metric. Additionally, we will report whether the fraction of latrines for which we use motion sensors changes over time and whether any change differs between treatment and control groups. If so, we will report results with and without controls for the type of sensor used.

We may carry out other robustness tests. If so, we will clearly identify which tests were pre-specified and which tests were added ex post.

References

Anderson, M. L. (2008): "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American Statistical Association*, 103, 1481–1495.

Banerjee, A., R. Banerji, E. Duflo, R. Glennerster, and S. Khemani (2010a): "Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation in Education in India," *American Economic Journal: Economic Policy*, 2.

——— (2010b): "Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation in Education in India," *American Economic Journal: Economic Policy*, 2, 1–30.

Björkman, M., D. de Walque, and J. Svensson (2017): "Information is Power: Experimental Evidence on the Long-Run Impact of Community Based Monitoring," *American Economic Journal: Applied Economics*, 9.

Burlig, F., L. Preonas, and M. Woerman (2017): "Panel Data and Experimental Design," Working paper.

Clasen, T., D. Fabini, S. Boisson, J. Taneja, J. Song, E. Aichinger, A. Bui, S. Dadashi, W.-P. Schmidt, Z. Burt, and K. L. Nelson (2012): "Making Sanitation Count: Developing and Testing a Device for Assessing Latrine Use in Low-Income Settings," *Environmental Science & Technology*, 46.

Djimeu, E. W. and D.-G. Houndolo (2016): *Power calculation for causal inference in social science: Sample size and minimum detectable effect determination. 3ie Working Paper 26.*, New Delhi: International Initiative for Impact Evaluation (3ie).

Gibbons, C. E., J. C. S. Serrato, and M. B. Urbancic (2019): "Broken or Fixed Effects?" *Journal of Econometric Methods*, 8.

Imbens, G. W. and D. B. Rubin (2015): *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*, Cambridge University Press.

Lin, W. (2013): "Agnostic notes on regression adjustments to experimental data: Reexamining Freedman's critique," *Annals of Applied Statistics*, 7.

McKenzie, D. (2012): "Beyond baseline and follow-up: The case for more T in experiments," *Journal of Development Economics*, 99.

Reinikka, R. and J. Svensson (2005): "Fighting corruption to improve schooling: Evidence from a newspaper campaign in Uganda," *Journal of European Economic Association*, 3.

UNDP (2011): *Fighting Corruption in the water Sector: Methods, tools and good Practices*, United Nations Development Programme.

UNESCO (2014): "Puberty Education and Menstrual Hygiene Management," *Good Policy and Practice in Health Education - Booklet 9*, 1–59.

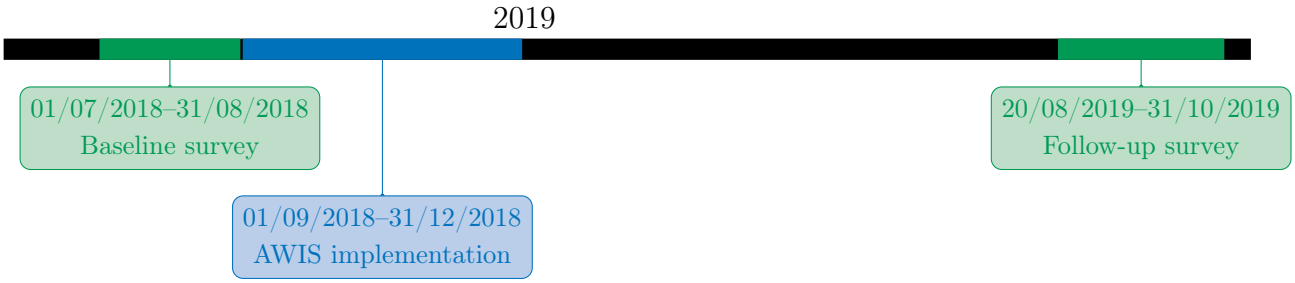
UNICEF (2012): *Water, Sanitation and Hygiene (WASH) in Schools: A companion to the Child Friendly Schools Manual*, UNICEF Division of Communication, New York.

Water Aid (2016): "School WASH research: Bangladesh country report," Report prepared for Water Aid by Jacques-Edouard Tiberghien, Partnerships in Practice.

Westfall, P. H. and S. S. Young (1993): *Resampling-based multiple testing: Examples and methods for p-value adjustment*, Wiley.

9 Figures

Figure 1: Project timeline



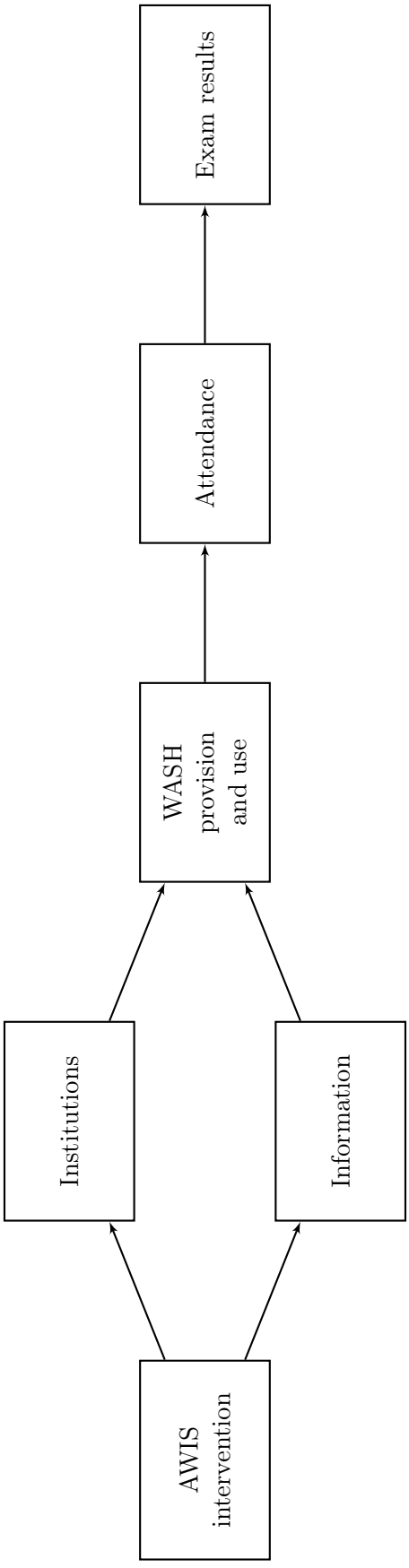


Figure 2: Theory of change

10 Tables

Table 1: Participants in AWIS workshops

	No. of participants /workshop	Share of participants
Parents	4	0.20
Students	4	0.20
SMC members	3.8	0.19
Teachers	3.6	0.18
Social leaders	3.6	0.18
Headteachers	0.4	0.02
Parents & SMC members	0.2	0.01
Local government/administrators	0.2	0.01
Teachers & SMC members	0.1	0.01

Table 2: AWIS Scores

	Mean	s.e.
Quality of sanitation facilities (1.82)		
Does the school have clear directives on facilities cleanliness and availability of soap, water and light?	1.54	0.02
Are responsibilities regarding the cleaning of sanitation facilities clear?	1.87	0.03
Are the toilets kept clean?	1.67	0.02
Are sanitation facilities well ventilated and with sufficient light?	1.70	0.03
Is sufficient water and soap available?	1.66	0.03
Are toilets accessible during school hours?	2.55	0.03
Does the school share information on water & sanitation management with parents during meetings?	1.49	0.03
Are parents and students able to influence the cleanliness and quality of sanitation facilities?	2.07	0.03
Gender (2.08)		
Does the school have clear directives on differentiated toilets for girls?	1.80	0.03
Are there separated toilets for boys and girls?	2.28	0.04
Do all girls use the toilets?	2.38	0.03
Are girls and their parents able to influence the suitability of toilets to the needs of girls?	1.85	0.03
Menstrual Hygiene (1.48)		
Does the school have clear directives on MHM?	1.27	0.02
Are girls talked to about MHM and are sanitary napkin made available to them?	1.57	0.03
Is there a sanitation block for girls to take care of their menstrual hygiene?	1.29	0.02
Are there any teachers responsible to teach girl students on MHM?	1.44	0.03
Are girls and parents able to influence the quality of support on MHM at school?	1.84	0.03
Disabled (1.25)		
Does the school have clear directives on the adaptation of sanitation facilities for disabled students?	1.17	0.02
Are there toilets adapted for disabled children?	1.07	0.01
Are disabled children and their parents able to influence the suitability of toilets to their needs?	1.52	0.03
Budget and expenditure (1.80)		
Does the school have clear financial planning?	1.59	0.03
Does the school have clear procedures for accounting?	1.99	0.03
Are financial planning procedures well implemented?	1.53	0.02
Are accounting procedures well implemented?	1.67	0.02
Are financial audits implemented?	1.90	0.03
Are parents, teachers and the SMC able to influence financial planning and management?	2.09	0.03

Notes: Statistics averaged across participants.

Table 3: Top priorities identified during the AWIS workshop

Priority area for action	Fraction of schools
Construct toilet cubicles	0.80
Water and soap should be available inside toilet cubicles	0.80
Improve cleanliness	0.73
General improvements to WASH facilities	0.57
Wash blocks for girls/MHM friendly toilets	0.50
Separate toilets for boys and girls	0.50
Provision of sanitary napkins	0.50
Provision of safe drinking water	0.47
Increase awareness among students on hygiene and MHM practices	0.47
Construct toilets for disabled students	0.43
Increase awareness	0.37
Increase the school budget for sanitation facilities	0.37
Hire cleaning staff	0.33
Appoint female teacher to discuss about MHM	0.33
Covered bin should be available inside toilet cubicles	0.30

Notes: Table lists priority areas for action identified by participants at workshop. Fraction of schools refers to the fraction of schools for which workshop participants listed the given area among the priority areas at the workshop.

Table 4: Most common actions identified during the AWIS workshop

Action	Fraction of schools
Contact local politicians/officials (eg. Union Parishad, upazila, local MP, etc)	1.00
Meetings between parents, teachers, head teacher and SMC	0.90
Increase awareness among students	0.83
Students take responsibility of cleaning	0.73
Raise funds from social leaders, wealthy individuals, parents	0.63
Voluntary contribution of workshop participants	0.53
Increase monitoring	0.47
Allocate separate budget/increase budget to sanitation facilities	0.47

Notes: Table lists possible actions identified by participants at workshop. Fraction of schools refers to the fraction of schools for which workshop participants listed the given action among possible actions at the workshop.

Table 5: Strengths of the AWIS workshop

Strength	Fraction of schools
Learning about WASH safe standards and practices	0.97
Important to discuss about school water and sanitation facilities	0.90
Presentation and facilitation by project staff	0.90
Discussion and increased awareness of MHM safe standards and practices	0.83
Participatory, systematic and constructive discussion	0.73
Representation of different categories of school stakeholders	0.43
Overall workshop logistics and management	0.33
Discussion of problems and possible solutions	0.33

Notes: Table lists perceived strengths of the AWIS workshop as reported by workshop participants. Fraction of schools refers to the fraction of schools for which workshop participants listed the given strength among the strengths of the AWIS workshop.

Table 6: Weaknesses of the AWIS workshop

Weakness	Fraction of schools
Time for the workshop was too short	1.00
No use of multimedia, case studies or practical examples	0.97
Need to involve more/all students	0.87
Isolated initiative, similar activities are needed more often/regularly	0.80
Venue and logistics (food provided, afternoon/morning, facilitators, etc)	0.53
No budget to cover transportation costs for participants	0.43
Need to involve more participants	0.37
Need long-term support from NGOs for monitoring and future activities	0.33
Limited possibilities to implement decisions and actions taken during the workshop	0.30

Notes: Table lists perceived weaknesses of the AWIS workshop as reported by workshop participants. Fraction of schools refers to the fraction of schools for which workshop participants listed the given weakness among the weaknesses of the AWIS workshop.

Table 7: Balance of school characteristics

	Control	Treated	Obs
Number of students	503.2	648.4	60
	(56.56)	(78.06)	
Number of teachers	12.4	12.5	60
	(0.43)	(0.55)	
Student/teacher ratio	40.4	52.7*	60
	(4.01)	(5.89)	
Mixed gender school	0.80	0.83	60
	(0.07)	(0.07)	
Madrasah	0.27	0.33	60
	(0.08)	(0.09)	
Distance from district capital	4.9	5.8	60
	(0.81)	(0.92)	
Number of buildings	3.0	3.0	60
	(0.16)	(0.25)	
Years since principal appointment	19.2	16.5	60
	(1.71)	(1.86)	
Principal education: MA or higher	0.63	0.73	60
	(0.09)	(0.08)	
Number of students group and parent/teacher organizations	1.7	1.9	60
	(0.18)	(0.17)	
The school has a PTA	0.80	.97*	60
	(0.07)	(0.04)	
Share of female SMC members	0.19	0.17	58
	(0.01)	(0.01)	
Budget exists for the ay 2017-2018	0.53	0.68	58
	(0.09)	(0.09)	
Number of programs received by the school in the ay 2017-2018	1.6	.9**	59
	(0.23)	(0.13)	
Hygiene is taught in school	0.50	0.50	60
	(0.09)	(0.09)	
MHM is taught in school	0.13	0.27	60
	(0.06)	(0.08)	
Source of drinking water is bacteria contaminated	0.47	0.52	59
	(0.09)	(0.10)	
Share of teachers' toilets	0.12	.22**	60
	(0.03)	(0.03)	
Pvalue of F-test for joint significance		0.332	
Pvalue of Hotelling's T-Squared test		0.194	

Notes: The table reports means and standard errors (in parentheses) of baseline characteristics in treatment and control groups. Regressions are run at school level with upazila FE. Significance levels are obtained from randomization based inference, reassigning treament status at the community level with 500 repetitions. p-values from joint F test and Hotelling's T-squared test joint significance of differences on all listed variables between treatment and control groups. *** p <0.01, ** p <0.05, * p <0.1.

Table 8: Balance of toilet characteristics

	Control	Treated	Obs
Years since toilet construction	7.9	8.2	242
	(1.21)	(1.15)	
Toilet accessible to disabled students	0.087	0.054	242
	(0.04)	(0.04)	
Toilet can be closed	0.83	0.82	242
	(0.05)	(0.06)	
Toilet is functioning	0.79	0.85	242
	(0.04)	(0.05)	
Toilet is clean	0.51	0.46	242
	(0.08)	(0.07)	
Toilet is separated by gender	0.32	0.16	242
	(0.07)	(0.07)	
Handwashing station within same sanitation facility	0.22	0.24	242
	(0.06)	(0.07)	
Basic facility for anal cleansing	0.85	0.88	201
	(0.04)	(0.03)	
Improved facility for anal cleansing	0.66	0.69	201
	(0.06)	(0.05)	
Soap available	0.24	0.24	242
	(0.07)	(0.07)	
Cleaning products inside	0.34	0.27	242
	(0.07)	(0.07)	
Waste bin inside the toilet compartment	0.14	0.24	242
	(0.05)	(0.07)	
Pvalue of F-test for joint significance		0.590	
Pvalue of Hotelling's T-Squared test		0.778	

Notes: The table reports means and standard errors (in parentheses) of baseline characteristics in treatment and control groups. Regressions are run at toilet level with upazila FE and weighted by the number of toilets/student toilets per school. Standard errors are clustered at school level. Significance levels are obtained from randomization based inference, reassigning treament status at the community level with 500 repetitions. p-values from joint F test and Hotelling's T-squared test joint significance of differences on all listed variables between treatment and control groups. *** p <0.01, ** p <0.05, * p <0.1.

Table 9: Sampling strategies: survey respondents

Category	Sampling
Headteacher	**Baseline and follow-up** We interview all headteachers
Teaching staff	**Baseline** Within each school, we randomly selected one male and one female teacher for interview from the full roster of teaching staff present during our visit. In addition, we interviewed all teachers with special responsibility for WASH and/or MHM e.g. physical education teachers or teachers responsible for classes for girls on MHM.
	Follow-up Within each school, we construct a random sample of teachers which includes: at least one male and one female teacher selected at random from the full roster; at least one male and one female teacher with special responsibility for WASH and/or MHM, if there are any in the school; at least one male and one female teacher who participated in the WIN workshop, if there are any; and at least one male and one female teacher who did not participate in the WIN workshop, if there are any.[30]
Students	**Baseline** Within each school, we select one female and one male student per school for the interview. In order to interview students who are well-informed and will feel comfortable to conduct the interview, we do not select students randomly. Instead, we select "class captains" from the highest available grade. "Class captain" is a role assigned by teachers to the most responsible and best performing students. If there is more than one class captain in the highest grade of either gender, we randomly select between them.[31]
	Follow-up We adopt the same sampling strategy as at baseline. In addition, if the students sampled using this approach did not participate in the AWIS workshop, we additionally ask headteachers to nominate a male and/or female student among the AWIS participants, if present at the school on the day of the interview.
Cleaning staff	**Baseline and follow-up** We interview at least one member of the cleaning staff if the school has cleaning staff who are present during the interview. If the school has more than one janitor, we randomly select one to interview.

[30] In practice, we achieve this by sampling with replacement, for each gender, one teacher from the full sample, one teacher from those with special responsibilities, one teacher which attended WIN and one which did not. Teachers may be sampled under multiple criteria. The number of teachers interviewed will range between 1 and 4.

[31] If there were no class captains, we specified that we would survey students from higher grades that have held some equivalent responsible role in the school e.g. sport captains or mentors to younger students. In practice, we identified class captains in all schools at baseline.

Table 10: Balance of main outcomes

	Control	Treated	Obs
A) Knowledge			
Head-teacher knowledge	0.00	-0.034	60
	(0.11)	(0.10)	
Teachers knowledge	0.00	0.093	176
	(0.07)	(0.08)	
Students knowledge	0.00	0.0043	108
	(0.08)	(0.09)	
B) Processes and institutions			
Quality of institutional processes	-0.0025	0.0056	60
	(0.07)	(0.06)	
C) WASH provision and use			
Number of student toilets	1.5	1.7	58
	(0.09)	(0.11)	
Student toilet quality index	-0.0035	-0.046	242
	(0.05)	(0.05)	
Daily toilet events	8.5	11.1	191
	(1.58)	(1.86)	
D) Attendance and exam results			
Daily attendance	0.65	0.63	885312
	(0.02)	(0.02)	
Exam results	2.3	2.3	10717
	(0.08)	(0.04)	

Notes: Standard errors clustered at school level. Upazila FE are included in all regressions. The test on "Number of student toilets" uses regressions at school level. The test on "Student toilet quality index" uses regressions at toilet level, weighted by the number of student toilets per school. The test on "Daily toilet events" uses regressions at toilet level, with magnetic/motion sensor FE and weighted by the number of installed sensors and student toilets per school. The tests on attendance use regressions at student level, with month and grade FE and weighted by the number of students per grade-month. The tests on exam results use regressions at student level, with JSC/SSC exam FE and weighted by the number of student writing that exam per school. *** p <0.01, ** p <0.05, * p <0.1.

Table 11: Regression specification

	Outcome	Unit of observation	Time fixed effects	School fixed effects
A1	Head-teacher knowledge	School	Upazila-year	School
A2	Teacher knowledge	Teacher	Upazila-year	School
A3	Student knowledge	Student	Upazila-year	School
A4	Quality of institutional processes	School	Upazila-year	School
B1	Number of functioning toilets	School	Upazila-year	School
B2	Toilet quality index	Toilet	Upazila-year	School
B3	Latrine use	Toilet	Upazila-year	School
C1	School attendance	Student-day	Day-grade-upazila-year	School-grade-day
C2	Exam results	Student	Level-upazila-year	School-level

Notes: Day fixed effects refer to Julian day fixed effects. Year fixed effects refer to calendar year fixed effects, corresponding to periods pre and post intervention.

Appendices
Supplementary material

A Informed consent/assent

Eligible schools are identified by our implementing partner. Recruitment is finalized during school visits, as we obtain informed oral consent from the headteacher to conduct the study in the school. We obtain consent from the headteacher separately for the surveys and intervention. Enumerators explain the purpose of the study and the structure of the data collection before obtaining informed consent. The different parts of the study then have different consent procedures.

- **Interviews with staff** We collect individual informed consent of each adult interview subject individually. Critically, however, we do not ask questions which relate to private individual information, but instead ask questions regarding general practice and collective experience.
- **Interviews with students** The headteacher, acting in loco parentis, consents to the process of interviewing students. Students then give individual assent to participate in the interviews. Critically, however, we do not ask questions which relate to private individual information, but instead ask questions regarding general practice and collective experience.
- **Direct observations of WASH facilities** These are carried out with the headteacher's consent. However, they involve no interaction with any individual, nor any individually identifiable information. As a result they require no further consent.
- **Latrine use intensity data** These data are collected with the headteacher's consent. However, by design, they do not collect individually identifiable information, instead collecting only anonymous data on whether or not the latrine is in use (via motion sensors) or recording when doors are opened and closed (door sensors). As a result, they require no further consent process.
- **Attendance data** These data are collected with the headteacher's consent. Since these data only involve normal educational practice, we do not collect individual informed consent to collect these data.
- **Exam result data** These data are publicly available. We thus do not collect consent to examine these data.

Whenever we collect consent or assent, and in particular before the start of each interview with an individual subject, enumerators explain that participation in the survey is voluntary and not compulsory, that they can withdraw their participation at any time, and that they can decide to not answer to any of our questions. We explain that participation of the school in the study or the intervention is not conditional on respondents completion of the interview, nor on their specific answers to our questions.

We obtain informed oral consent or assent, as appropriate, from each interviewee individually (Appendix A.1, A.2). Given that the research involves minimal risks and does not collect data of a private, sensitive, incriminating or psychologically distressing nature, the research team collects only oral consent.

All recruitment and consent procedures and study materials are administered in Bengali. Informed oral consent is obtained in Bengali.

We do not disburse any compensation to any study subjects.

A.1 Informed consent/assent from headteachers

"I am working for a NGO called NGO Forum for Public Health, and collaborating with researchers from Stockholm University. We are working on a research project in this region to assess the status of water and sanitation facility in schools in this region. As you may know, this school is participating in this project.

As part of the research project, we want to collect some observational data in the school. We also want to ask some questions about the school's water and sanitation facilities to you, few teachers, the cleaning staff (if any) and few students. The questions we ask are not personal, but are about general issues facing all students and staff.

We will take very good care of all data we will collect and no one who is not connected with the project will have access to school information or to the personal information (like names) of people interviewed. We will only use personal information, like names, in carrying out this project, and if we use information from the survey in the future we will remove names and change locations so that no one can recognize the school and its staff. There are no other risks to participating in this project.

The full survey will take few hours. The personal interview with each school staff will take about 30 minutes each. Respondents do not have answer any questions if they do not want to. If respondents agree to listen to our questions, they can choose not to answer any question, and they can stop answering questions at any time. There are no penalties or loss of benefits for the school if respondents don't answer any question, or if they ask us to stop asking questions. There are no other costs involved to answering our questions, beyond the time it will take to do so. We will do our best to minimize disruption to school activities while we carry out the survey. There is no personal benefit to respondents of answering our questions. However, answering the questions could help us and other NGOs to design better projects in the future to provide better water and sanitation facilities in schools like this one.

If you have any other questions about this project, or would like to offer any input, please contact Ahasan Habib at NGO Forum on 019172322xx.

Do you agree to participate in the interview?"

A.2 Informed consent/assent from teachers, students, cleaning staff and School Management Committee members

"I am working for a NGO called NGO Forum for Public Health, and collaborating with researchers from Stockholm University. We are working on a research project in this region to assess the status of water and sanitation facility in schools in this region. As you may know, this school is participating in this project.

As part of the project, we would like to ask you few questions. We will take very good care of your data and no one who is not connected with the project will have access to your personal information, like your name. We will only use your personal information, like names, in carrying out this project, and if we use information from the survey in the future we will remove names and change locations so that no one can recognize you. There are no other risks to participating in this project.

The interview will take about 30 minutes. You do not have to answer any questions if you do not want to. If you agree to listen to our questions, you can choose not to answer any question, and you can stop answering questions at any time. There are no penalties or loss of benefits for the school or yourself if you don't answer any question, or if you ask us to stop asking questions. There are no other costs involved to answering our questions, beyond the time it will take to do so. There is no benefits for respondents of answering our questions. However, answering the questions could help us and other NGOs to design better projects in the future to provide better water and sanitation facilities in schools like this one.

If you have any other questions about this project, or would like to offer any input, please contact Ahasan Habib at NGO Forum on 019172322xx.

Do you agree to participate in the interview?"

B AWIS - School WASH Scorecard

B.1 Quality of sanitation facilities

- Does the school have clear directives on sanitation facilities cleanliness and availability of soap, water and light?

    1. No, or directives are not written
    2. Yes, there are written directives but they could be improved
    3. Yes, clear and good written directives are available

- Are responsibilities regarding the cleaning of sanitation facilities clear?

    1. Responsibilities for cleaning are not clear
    2. Yes, there is a cleaning person but responsibilities beyond (for example of the school management committee and users) are not clear, or responsible people need clearer instructions
    3. Yes, the responsibilities of the School Management Committee the cleaning person and the users are clear

- Are the toilets kept clean?

    1. No, they are quite dirty most of the time
    2. Mostly yes, but cleaning could be improved
    3. Yes, a responsible person, supported by teachers and students, keeps them continuously clean

- Are sanitation facilities well ventilated and with sufficient light?
    1. Frequently sanitation facilities are not well ventilated, they have little light.
    2. Mostly yes, but could be improve (e.g. some element is missing or availability is intermittent)
    3. Yes, sufficient ventilation and light are always available
- Is sufficient water and soap available?
    1. Frequently sanitation facilities have no soap or sufficient water are available
    2. Mostly yes, but could be improve (e.g. some element is missing or availability is intermittent)
    3. Yes, sufficient water and soap are always available
- Are toilets accessible during school hours?
    1. No, toilets are frequently locked during school hours
    2. Mostly yes, but toilets may be sometimes locked during school hours
    3. Yes, toilets are always accessible for students and teachers during school hours
- Does the school share information related to water and sanitation management with parents during meetings?
    1. No information on water and sanitation management is shared with parents during meetings
    2. Some information on water and sanitation management is shared with parents during meetings, but not always and not in a complete and clear way
    3. Yes, the school updates parents on the status of water and sanitation facilities in every meeting. And the information shared is clear and complete
- Are parents and students able to influence the cleanliness and quality of sanitation facilities?
    1. The school considers that parents and students do not need to be involved in the evaluation of the cleanliness and quality of sanitation facilities
    2. The school allows parents and students to participate in the evaluation of the cleanliness and quality of sanitation facilities, but no effective complaint mechanisms are available
    3. Parents and students can voice complaints and these are analysed and addressed

B.2 Gender

- Does the school have clear directives on differentiated toilets for girls?
    1. No, or directives are not written
    2. Yes, written directives indicate that toilets for girls must be separated but do not obliged to have a plastic bin with cover available in girl's toilets
    3. Yes, written directives indicate that toilets for girls must be separated and a plastic bin with cover must be available in girl's toilets
- Are there separated toilets for boys and girls?
    1. No, there are only mix toilets
    2. Yes, but gender separation is not always respected
    3. Yes, clearly separated
- Do all girls use the toilets?
    1. No girls uses the toilets
    2. Some girls use the toilets
    3. All girls use the toilets
- Are girls and their parents able to influence the suitability of toilets to the needs of girls?
    1. The school considers that girls and their parents do not need to be involved in the evaluation of the suitability of toilets to the needs of girls
    2. The school allows girls and their parents to participate in the evaluation of the suitability of toilets to the needs of girls, but no effective complaint mechanisms are available
    3. Girls and their parents can voice complaints and these are analysed and addressed

B.3 Menstrual Hygiene

- Does the school have clear directives on menstrual hygiene management?
    1. No, or directives are not written
    2. Yes, there are written directives but they could be improved
    3. Yes, clear and good written directives are available
- Are girls talked to about menstrual hygiene management and are sanitary napkin made available to them?
    1. No, nobody talks to the girls about menstrual hygiene management
    2. Yes, but could be improved

3. Yes, a teacher talks about menstrual hygiene management with the girls and makes sure they get sanitary napkins

- Is there a sanitation block for girls to take care of their menstrual hygiene?

  1. No
  2. Yes, but not properly maintained
  3. Yes and properly maintained

- Are there any teachers responsible to teach girl students on menstrual hygiene management?

  1. No, nobody is assigned
  2. Yes, a female teacher is assigned but not monitored
  3. Yes and properly monitored

- Are girls and their parents able to influence the quality of support on menstrual hygiene management at the school?

  1. The school considers that girls and their parents do not need to be involved in the evaluation of the quality of the support on menstrual hygiene management at the school
  2. The school allows girls and their parents to participate in the evaluation of the quality of the support on menstrual hygiene management at the school, but no effective complaint mechanisms are available
  3. Girls and their parents can voice complaints and these are analysed and addressed

B.4 Disabled

- Does the school have clear directives on the adaptation of sanitation facilities for disabled students?

  1. No, or directives are not written
  2. Yes, there are written directives but they could be improved
  3. Yes, clear and good written directives are available

- Are there toilets adapted for disabled children?

  1. No, they are not adapted
  2. Yes, but adaptation could be improved
  3. Yes, they are completely adapted

- Are disabled children and their parents able to influence the suitability of toilets to their needs?

1. The school considers that disabled children and their parents do not need to be involved in the evaluation of the suitability of toilets to their needs
2. The schools allows disabled children and their parents to participate in the evaluation of the suitability of toilets to their needs, but no effective complaint mechanisms are available
3. Disabled children and their parents can voice complaints and these are analysed and addressed

B.5 Budget and expenditure

- Does the school have clear financial planning?

  1. No clear financial planning, or planning not written
  2. Yes, there is a written financial planning but it could be improved
  3. Yes, clear and updated financial planning is available

- Does the school have clear procedures for accounting?

  1. No procedures for accounting exist, or they are not written
  2. Yes, there are written procedures for accounting but the accounting system could be improved
  3. Yes, a clear and reliable accounting system is available and updated

- Are financial planning procedures well implemented?

  1. No, financial planning procedures are mostly not applied
  2. Yes, but with limitations. For example the budget is rarely monitored or observed
  3. Yes, financial planning procedures are carefully followed and the budget is observed

- Are accounting procedures well implemented?

  1. No, accountability is not done reliably (e.g. manually)
  2. Yes, but with limitations. For example accounting is computerized, but the system could be improved
  3. Yes, accounting procedures are carefully followed and the system used is safe from failures and tampering

- Are financial audits implemented?

  1. No, or very limited and/or unreliable information
  2. Yes, but could be improved
  3. Yes, clear, complete and reliable financial audits are implemented

- Are parents, teachers and the school management committee able to influence financial planning and management?

    1. Parents, teachers and the schools management committee are not able to evaluate financial planning and management in the school
    2. Parents, teachers and the schools management committee can identify problems regarding financial planning and management in the school, but there is no effective complaint mechanism
    3. Parents, teachers and the schools management committee can voice complaints and these are analysed and addressed

C Construction of knowledge indices

A1, A2: Head-teacher knowledge and teacher knowledge

1. Share of correct responses given when asked about the characteristics that school WASH facilities should have in order to respect hygiene and safety standards and guarantee equal access to boys and girls. The correct responses are:

   (a) There should be an adequate number of sources of drinking water;

   (b) Sources of drinking water and their surroundings should be clean to avoid contamination of pollutants;

   (c) Sources of drinking water and their surrounding should be accessible for all, including disabled students;

   (d) There should be an adequate number of toilets in the school;

   (e) Toilets should be cleaned regularly;

   (f) Toilets should be accessible for all, including disabled students;

   (g) Soap should be available in the sanitation facility;

   (h) In the vicinity of toilets there should be water points (sink or tubewells) for washing hands;

   (i) Toilets should be big enough to guarantee ventilation;

   (j) The toilets should not be dark (large windows or light bulb);

   (k) Toilets should be separated by gender.

2. Share of correct responses given when asked about good hygienic practices that students should follow before eating or after toilet use. The correct responses are:

   (a) Anal cleansing after defecation;

   (b) Wash hands with soap after toilet use;

   (c) Wash hands with soap before eating.

3. Share of correct responses given when asked about the characteristics that school sanitation facilities should have in order to guarantee that girls can use toilets during their periods and properly manage their periods in hygienic and safe conditions. The correct responses are:

   (a) Toilets should be separated by gender;

   (b) Girls bathrooms should be organized such that girls can change in privacy (example: doors, dedicated changing rooms, etc);

   (c) Girls toilets should have a mechanism for the collection and disposal of sanitary protection materials, such as in a pit or incinerator or plastic bin;

   (d) The bin in girls-toilets for sanitary pads disposal should be covered and it should be easy to clean;

(e) It should be possible for girls to privately wash cloths used to absorb menstrual blood after changing, e.g. there should be a separate washing station or washing facility inside toilet cubicles used by girls when they have their period;

(f) There should be a small mirror (even a broken piece of mirror) provided in toilet cubicles to help girls check for spotting or leaking and ensure everything is in order before leaving.

4. Share of correct responses given when asked about good hygienic practices that girls should follow during their period. The correct responses are:

   (a) Use a clean material to absorb or collect menstrual blood (sanitary pad or clean cloth);

   (b) Use soap and water for washing the body as required;

   (c) Take enough food and water to ensure their body balance and strength.

A3: Student knowledge

1. Share of correct responses given when asked about the good and hygienic practices that students should follow before eating or after toilet use. The correct responses are as listed above.

2. Share of correct responses given when asked about good hygienic practices that girls should follow during their period. The correct responses are as listed above.

3. Indicator for whether or not the respondent knows what the menstrual cycle is, evaluated by enumerators based on respondents' description of the menstrual cycle.

4. Indicator for whether the respondent knows what a sanitary pad is, evaluate by enumerators based on on respondents' description of a sanitary pad.

D Construction of institutions index

A4: Quality of institutional processes

1. The share of necessary clear written directives that the school has regarding sanitation facilities. The list of necessary directives is:

   (a) Cleanliness of sanitation facilities;

   (b) Availability of soap in washrooms/toilets;

   (c) Availability of water in washrooms/toilets;

   (d) Availability of light in washrooms/toilets;

   (e) Gender-separated washrooms/toilets;

   (f) Availability of disposal mechanisms of sanitary pads;

   (g) Menstrual Health Management;

   (h) Accessibility to sanitation facilities for disabled students.

2. The ratio of teachers responsible to speak with girls about MHM to the number of female students in the school.

3. Frequency of parent meetings, measured by the share of months per year when a meeting occurs.

4. Whether the school submitted an official report on water and sanitation facilities in the school relative to the last academic year to the respective authority (the responsible Education Officer);

5. Whether the school had a budget for the last academic year;

6. Whether the school had a budget specifically allocated to i) construction, ii) repair, or iii) cleaning of WASH facilities.

E Construction of sanitation facility quality index

B1: Quality of sanitation facilities We construct the index for each toilet cubicle available to students. The sub-indices are:

1. Whether the toilet is functioning;
2. Whether the toilet cubicle has access for people with a physical disability (including wheelchair users);
3. Whether the toilet cubicle has a door that can be closed properly;
4. Whether the toilet is clean;
5. Whether the toilet is separated by gender;
6. Whether there is a handwashing station within the same sanitation facility;
7. Whether there is any facility for anal cleansing inside the toilet cubicle, defined as toilet paper, running water from a functioning tap/spray, water stored inside the cubicle, or a toilet kettle to collect water from outside the toilet cubicle;
8. Whether there is an improved facility for anal cleanings inside the toilet cubicle, defined as running water from a functioning tap/spray or water stored inside the cubicle.
9. Whether there is soap inside the same sanitation facility;
10. Whether there are cleaning materials available inside the toilet cubicle;
11. Whether there is a waste bin inside this toilet cubicle.

F Construction of toilet use event data

Door sensors The magnetic sensors we install on latrine cubicle doors consist of two parts, one of which consists of a reed switch and the second of which consists of a magnet. When the magnet is less than 13mm away from the reed switch, the switch closes. When connected to the event logger, the raw data output consists of a series of events which are coded zero when the current switches off, implying that the magnet moves more than 13mm away from the switch, which it does when the door opens, and coded one when the current switches on, implying that the magnet moves to within 13mm away from the switch, which it does when the door closes.

We identify potential latrine use events by looking for periods of between 30 seconds and 10 minutes when the door is closed, framed by periods when the door is open on either side. These correspond to potential use events. The classification is not sensitive to whether or not the door is habitually left open or closed between switches. Our field staff report that both are common, as some doors do not have a lock or other closing device on the outside of the door.

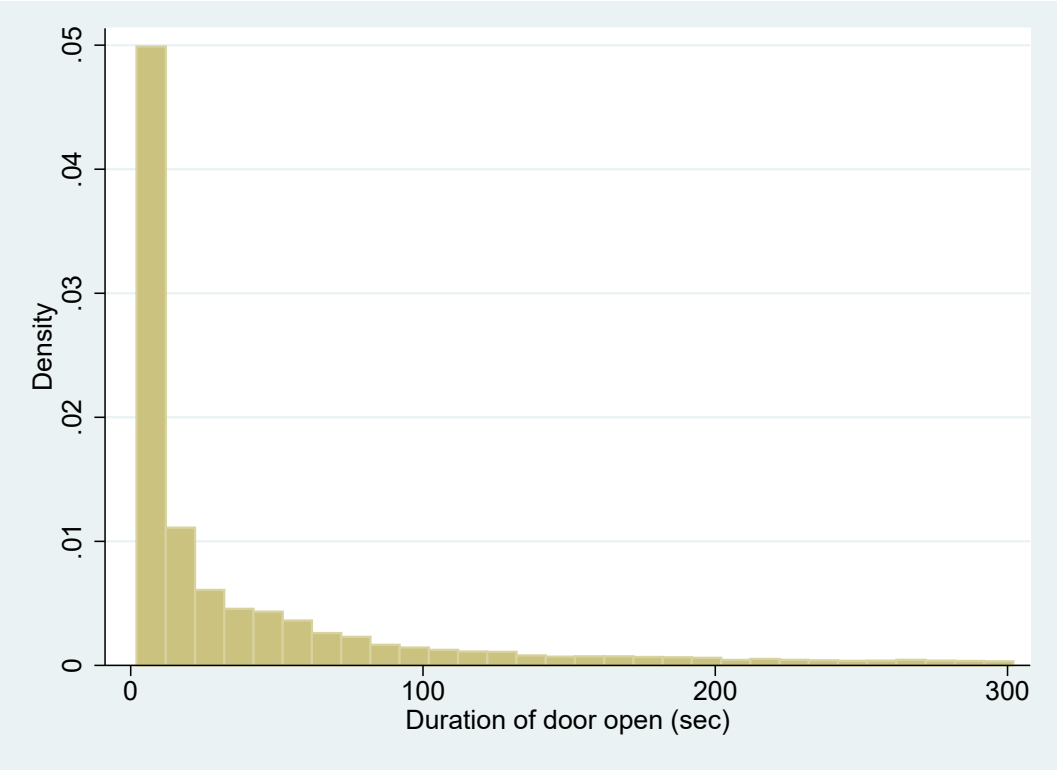
Our classification of potential latrine use events may overcount latrine use events during periods of relatively intensive use. This is because our classification will count any period of between 30 seconds and 10 minutes when the door is closed as a potential latrine use event. If the door is closed for a period within this range between two consecutive uses, we will classify this as a potential latrine use event.

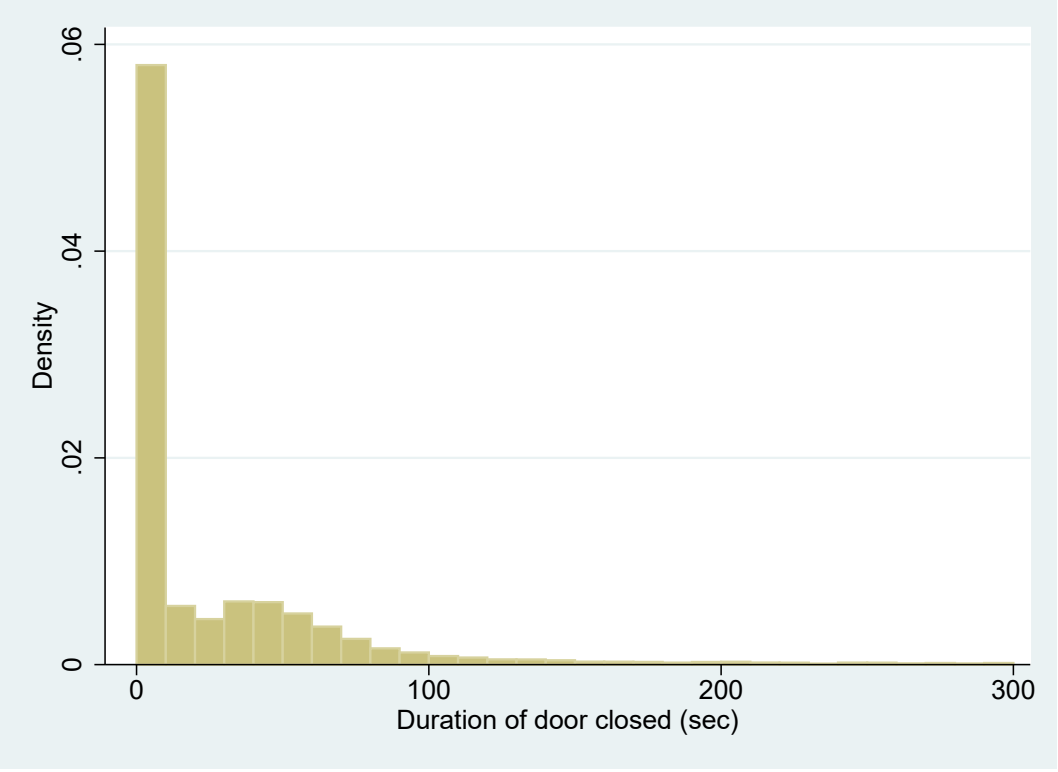
Before identifying potential latrine use events, we clean the data by dropping all very short (< 2 seconds) openings of the door. We remove these events because occasionally the current breaks briefly, creating short opening events in the data which are too short to correspond to actual openings. We choose the threshold of 2 seconds because we estimate that it takes a minimum of 3 seconds to open a door, walk through it, and close it again. We use the cleaned data to classify potential latrine use events.

Appendix Figure F1 shows the processed data, focusing on events shorter than 300 seconds. The top panel shows a histogram of the durations of periods when the door is open, which varies smoothly across event durations. The bottom panel shows a histogram of the durations of periods when the door is closed. The hump in the distribution which peaks around 50 seconds corresponds to potential latrine use events.

The classification is insensitive to what number we use as the upper threshold to define latrine use events, because there are few events with door closures near ten minutes.

Figure F1: Data from magnetic sensors





Motion sensors We install motion sensors when it is not possible to install door sensors, for example because the cubicle does not have a door or because the design of the door means that installing the magnetic sensor is not possible. The motion sensors we install are passive infra-red sensors. The infra-red sensors detect when a warm body moves in or out of its frame of reference. When they do so, the sensor emits a signal which is recorded by the data logger.

We construct a sequence of potential latrine use events closely following Clasen et al. (2012), who also use infra-red sensors to detect potential latrine use events. Clasen et al. (2012) report that the typical signature of a latrine use event is a period of movement, followed by a longer period with no movement triggers corresponding to the squatting period of the latrine use event, followed by another period of movement.

The algorithm proceeds in the following steps:

1. We group events into "edges" if they occur within 15 seconds of a previous event.

2. We drop edges of durations of less than 30 seconds with no neighbours within 3 minutes, as these most likely correspond to isolated triggers and not to latrine use.[32]

3. We group edges into activities when they occur within 10 minutes of the start of the most recent edge.

4. For activities longer than 6 minutes, we divide them into multiple activities if there are periods of 3 minutes or longer during which there are no edges, assuming that latrine users are unlikely to be still for periods without triggering the sensor for periods of longer than 3 minutes.

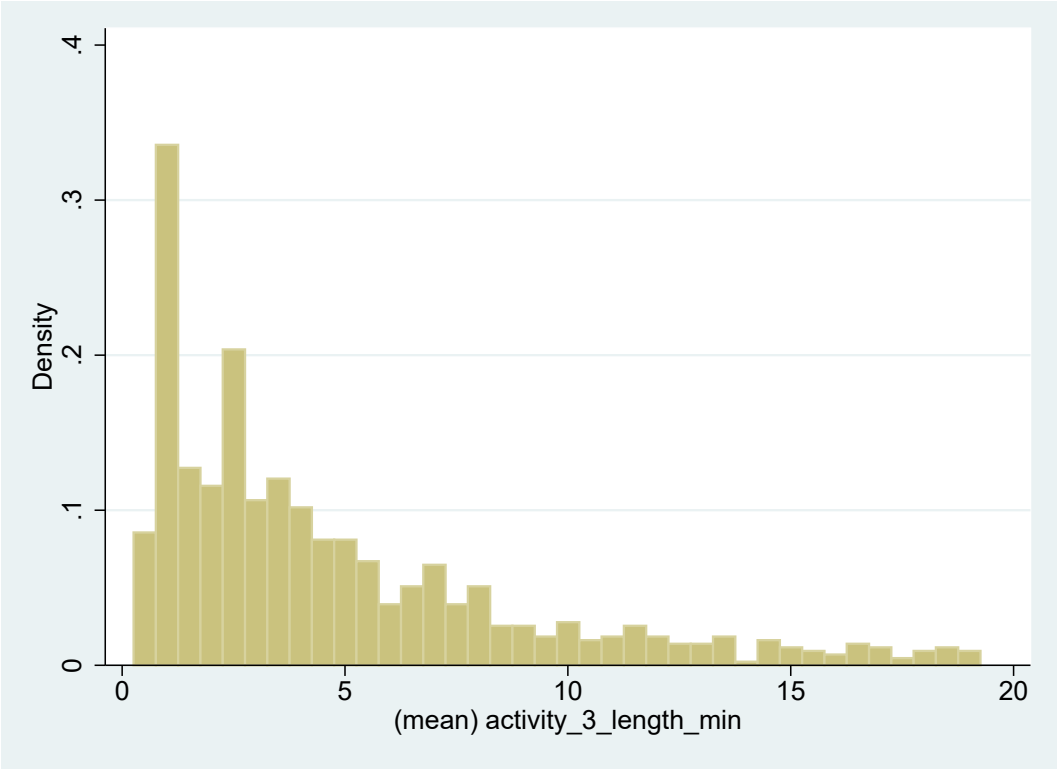
5. We classify the resulting set of activities as potential latrine use events.

In contrast to the magnetic door sensor data, the motion sensor data potentially undercounts use during periods of frequent use, because we classify long periods of events as a single activity as long as they are not broken by a period of at least 3 minutes of inactivity. Some very long events are identified in the data, suggesting that we may be undercounting events during these periods, although these periods could also correspond to cleaning events or other non-latrine use triggers

Appendix Figure F2 shows the distribution of event times for latrine use events obtained from the motion sensor data.

[32]Clasen et al. (2012) used ten minutes at this stage. When we used ten minutes, we discovered a large number of very short isolated events generated after dividing longer activities by periods of three minutes or more.

Figure F2: Data from motion sensors



G Lasso algorithm

In one of our robustness checks, we will estimate the main effects using control variables, using the Lasso algorithm to select the optimal set of controls (Section 8).

The list of controls included in the Lasso algorithm will be the following:

- Number of students
- Number of teachers
- Student/teacher ratio
- Mixed gender school
- Madrasah
- Distance from district capital
- Number of buildings
- Years since principal appointment
- Principal education: MA or higher
- Number of student groups and parent/teacher organizations
- The school has a PTA
- Share of female SMC members
- Number of programs received by the school in the ay 2017-2018
- Hygiene is taught in school
- MHM is taught in school
- Source of drinking water is bacteria contaminated
- Share of teachers' toilets

H Supplementary Tables and Figures

Table H1: Correlations between AWIS scores and survey observation

Dep. var.: AWIS scores	Toilets clean	Light and ventilation	Soap available	Toilets accessible	Toilets separated	Suitable for MHM	Accessible to disabled	Teacher responsible MHM
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Survey observation	0.58***	0.12	0.60***	-0.33	-1.28***	-0.03	0.09	2.28**
	(0.15)	(0.21)	(0.20)	(0.26)	(0.20)	(0.17)	(0.13)	(0.93)
N	592	592	590	587	575	571	563	571

Notes: Unit of observation is a scored given by an AWIS participant. Standard errors clustered by school. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table H2: Official and observed attendance

	Boys		Girls	
Dep. var.: Difference btw official and observed attendance	(3)	(4)	(5)	(6)
Grade 6	0.72	0.71	-0.35	-0.30
	(1.64)	(1.66)	(2.41)	(2.47)
Grade 7	2.86*	2.75*	0.98	0.97
	(1.55)	(1.46)	(1.40)	(1.39)
Grade 8	1.62	1.55	0.48	0.46
	(1.65)	(1.75)	(2.05)	(2.05)
Grade 9	0.03	0.01	1.30	1.29
	(0.99)	(1.01)	(1.00)	(0.95)
Grade 10	3.92***	3.94***	-0.55	-0.57
	(1.43)	(1.39)	(1.69)	(1.74)
School FE		✓		✓
N	179	179	242	242

Notes: The coefficients represent means of the difference between official and observed attendance, obtained from regressions at school-grade level with standard errors clustered at school level. When included, school FEs are demeaned. *** p <0.01, ** p <0.05, * p <0.1.

Table H3: Balance between completed and non-complete interviews at baseline

	Control	Treated	Obs
a) All teachers			
Interview not completed	0.340	0.305	282
	(0.04)	(0.04)	
Not contacted	0.183	0.143	282
	(0.04)	(0.03)	
Consent not given	0.157	0.162	282
	(0.03)	(0.03)	
a) Male teachers			
Interview not completed	0.242	0.287	188
	(0.04)	(0.04)	
Not contacted	0.0478	0.0528	188
	(0.02)	(0.03)	
Consent not given	0.194	0.234	188
	(0.04)	(0.04)	
a) Female teachers			
Interview not completed	0.358	0.276	94
	(0.07)	(0.06)	
Not contacted	0.358	0.276	94
	(0.07)	(0.06)	
Consent not given	0.00	0.00	94
	(.)	(.)	

Notes: Standard errors clustered at school level. Upazila FE are included in all regressions. *** p <0.01, ** p <0.05, * p <0.1.

Table H4: Correlates with interview non-completion at baseline

	Interview not completed			Not contacted			Consent not given		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Order of interview	0.04***	0.37***	0.06***	-0.02***	0.37***	-0.00	0.06***	0.00	0.07***
	(0.01)	(0.07)	(0.02)	(0.01)	(0.07)	(0.01)	(0.01)	(.)	(0.02)
Teacher responsible to teach WASH/MHM	0.08	-0.01	0.06	0.08	-0.01	0.09**	-0.00	0.00	-0.03
	(0.07)	(0.21)	(0.09)	(0.05)	(0.21)	(0.04)	(0.05)	(.)	(0.08)
Teacher gender	All	Female	Male	All	Female	Male	All	Female	Male
N	282	94	188	282	94	188	282	94	188

Notes: Standard errors clustered at school level. School FE are included in all regressions. *** p <0.01, ** p <0.05, * p <0.1.

Table H5: Balance: knowledge indicators

	Control	Treated	Obs
a) Head teachers			
Knowledge index	0.00	-0.034	60
	(0.11)	(0.10)	
Knowledge of WASH standard for school facilities	0.30	0.28	60
	(0.03)	(0.02)	
Knowledge of WASH practices	0.89	0.88	60
	(0.02)	(0.03)	
Knowledge of MHM standard for school facilities	0.29	0.28	60
	(0.03)	(0.02)	
Knowledge of MHM practices	0.62	0.70	60
	(0.04)	(0.03)	
a) Teachers			
Knowledge index	0.00	0.093	176
	(0.07)	(0.08)	
Knowledge of WASH standard for school facilities	0.23	0.26	176
	(0.01)	(0.01)	
Knowledge of WASH practices	0.83	0.82	176
	(0.02)	(0.02)	
Knowledge of MHM standard for school facilities	0.31	0.33	176
	(0.02)	(0.02)	
Knowledge of MHM practices	0.62	.69*	176
	(0.02)	(0.03)	
a) Students			
Knowledge index	0.00	0.0043	108
	(0.08)	(0.09)	
Knowledge of WASH practices	0.74	0.76	108
	(0.03)	(0.02)	
Knowledge of menstrual cycle	0.67	0.60	108
	(0.06)	(0.06)	
Knowledge of sanitary pad	0.70	0.70	108
	(0.06)	(0.06)	
Knowledge of MHM practices	0.63	0.60	108
	(0.03)	(0.05)	

Notes: Standard errors clustered at school level. Upazila FE are included in all regressions. *** p <0.01, ** p <0.05, * p <0.1.

Table H6: Balance: institutional processes quality

	Control	Treated	Obs
Quality of institutional processes	-0.0025	0.0056	60
	(0.07)	(0.06)	
School directives on WASH/MHM	0.096	0.096	60
	(0.03)	(0.02)	
Ratio MHM teachers to female students	0.0051	.0031*	60
	(0.00)	(0.00)	
Frequency of meetings with parents	0.46	0.50	55
	(0.06)	(0.06)	
Last a.y. report on WASH facilities submitted	0.60	0.60	60
	(0.09)	(0.09)	
Budget exists for the ay 2017-2018	0.53	0.68	58
	(0.09)	(0.09)	
Budget exists for construction, repair, or cleaning of WASH facilities for the a	0.23	0.36	58
	(0.08)	(0.09)	

Notes: Standard errors clustered at school level. Upazila FE are included in all regressions. *** p <0.01, ** p <0.05, * p <0.1.

Table H7: Balance: latrine use

	Control	Treated	Obs
a) All toilets			
Latrine use measured using motion sensor	0.25	.075*	128
	(0.09)	(0.03)	
Daily toilet events	8.5	11.1	191
	(1.58)	(1.86)	
Total duration of daily toilet events	23.1	20.3	191
	(6.60)	(3.80)	
Daily toilet events per student	0.057	0.068	59
	(0.01)	(0.01)	
Total duration of daily toilet events per student	0.14	0.13	59
	(0.04)	(0.03)	
a) Girls toilets			
Daily toilet events	9.3	13.7	100
	(2.19)	(2.86)	
Total duration of daily toilet events	17.3	19.9	100
	(4.11)	(4.44)	
Daily toilet events per student	0.067	0.086	59
	(0.02)	(0.02)	
Total duration of daily toilet events per student	0.15	0.14	59
	(0.04)	(0.03)	
a) Boys toilets			
Daily toilet events	6.3	7.2	73
	(2.66)	(1.68)	
Total duration of daily toilet events	25.3	20.7	73
	(16.88)	(6.35)	
Daily toilet events per student	0.076	0.052	48
	(0.02)	(0.01)	
Total duration of daily toilet events per student	0.25	0.14	48
	(0.10)	(0.04)	
a) Mixed toilets			
Daily toilet events	9.6	4.4	18
	(2.61)	(2.49)	
Total duration of daily toilet events	25.8	12.7	18
	(10.59)	(8.81)	

Notes: Standard errors clustered at school level. Regressions are at sensor or school level, with upazila FE and magnetic/motion sensor FE. *** p <0.01, ** p <0.05, * p <0.1.

Table H8: Balance: school attendance

	Control	Treated	Obs
Daily attendance	0.62	0.60	1103447
	(0.01)	(0.02)	
Daily attendance - boys	0.58	0.55	399582
	(0.02)	(0.02)	
Daily attendance - girls	0.65	0.63	699497
	(0.02)	(0.02)	

Notes: Standard errors clustered at school level. Regressions are at student level, with upazila, month and grade FE and weighted by the number of students per grade-month.*** p <0.01, ** p <0.05, * p <0.1.

Table H9: Balance: exam results

	Control	Treated	Obs
Exam results	2.3	2.3	10717
	(0.08)	(0.04)	
Exam passed	0.61	0.65	10717
	(0.02)	(0.01)	
Exam results - boys	2.2	2.4*	5131
	(0.08)	(0.08)	
Exam passed - boys	0.60	.67**	5131
	(0.02)	(0.02)	
Exam results - girls	2.3	2.3	5586
	(0.08)	(0.04)	
Exam passed - girls	0.63	0.61	5586
	(0.02)	(0.01)	

Notes: Standard errors clustered at school level. Regressions are at student level, with upazila FE and JSC/SSC exam FE and weighted by the number of student writing that exam per school. *** p <0.01, ** p <0.05, * p <0.1.